

SEQUENCE LISTING

<110> INVITROGEN CORPORATION
DALBY, Brian
BENNETT, Robert

<120> DELIVERY OF FUNCTIONAL PROTEIN SEQUENCES
BY TRANSLOCATING POLYPEPTIDES

<130> INVIT1280-1

<150> PCT/US00/08571

<151> 2000-03-31

<150> 60/127,467

<151> 1999-03-31

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 6404

<212> DNA

<213> Artificial Sequence

<220>

<223> vector pVP22/Myc-His

<400> 1

gacggatcgg	gagatctccc	gateccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttggtgtgt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttaggggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacgggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccgcg	tggctgaccg	cccaacgacc	360
cccgcctcatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggg	aaactgcca	cttggcagta	catcaagtgt	480
atcatatgcc	aagtagcgcc	cctattgacg	tcaatgacgg	taaattggcc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaattggcg	780
gtaggcgtgt	acgggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagt	900
taagcttatt	atgacctctc	gccgctccgt	gaagtcgggt	ccgcgggagg	ttccgcgcga	960
tgagtacgag	gatctgtact	acaccccgtc	ttcaggtatg	gcgagtcctg	atagtcgcgc	1020
tgacacctcc	cgccgtggcg	ccctacagac	acgctcgcgc	cagagggggc	aggctcggtt	1080
cgtccagtac	gacgagtcgg	attatgccct	ctacgggggc	tcgtcttcgc	aagacgacga	1140
acacccggag	gtcccccgga	cgcggcgtcc	cgtttccggg	gcggttttgt	ccggcccggg	1200
gcctgcgcgg	gcgcctccgc	cacccgctgg	gtccggaggg	gccggacgca	cacccaccac	1260
cgcccccccg	gccccccgaa	cccagcgggt	ggcgtctaag	gccccgcgg	ccccggcggc	1320
ggagaccacc	cgcggcagga	aatcggccca	gccagaatcc	gccgcactcc	cagacgcccc	1380
cgcgctcgacg	gcgccaaccc	gatccaagac	acccgcgcag	gggctggcca	gaaagctgca	1440
ctttagcacc	gcccccccaa	accccgacgc	gcatgggacc	ccccgggtgg	ccggctttaa	1500
caagcgcgtc	ttctgcgcgc	cggtcggggc	cctggcgccc	atgcatgccc	ggatggcggc	1560
tgtccagctc	tgggacatgt	cgcgtccgcg	cacagacgaa	gacctcaacg	aactccttgg	1620
catcaccacc	atccgcgtga	cggtctgcga	gggcaaaaac	ctgcttcagc	gcgccaacga	1680
gttggtgaat	ccagacgtgg	tgcaggacgt	cgacgcggcc	acggcgactc	gagggcgttc	1740

20/07/01 15:04:00

tgcggcgctcg	cgccccacccg	agcgacctcg	agccccagcc	cgctccgctt	ctcgccccag	1800
acggcccgc	gaggggtaccg	agctcggatc	cactagtcca	gtgtggtgga	attctgcaga	1860
tatccagcac	agtggcgggc	gctcgagctc	agagggcccc	cggttcgaac	aaaaactcat	1920
ctcagaagag	gatctgaata	tgcataccgg	tcatacatcac	catcaccatt	gagtttaaac	1980
ccgctgatca	gcctcgactg	tgccttctag	ttgccagcca	tctgttggtt	gccccctccc	2040
cgtgccttcc	ttgaccctgg	aaggtgccac	tcccactgtc	ctttcctaata	aaaatgagga	2100
aattgcatcg	cattgtctga	gtaggtgtca	ttctattctg	gggggtgggg	tggggcagga	2160
cagcaagggg	gaggattggg	aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	2220
ggcttctgag	gcggaaagaa	ccagctgggg	ctctaggggg	tatccccacg	cgccctgtag	2280
cggcgcatta	agcgcgggcg	gtgtggtggt	tacgcgcagc	gtgaccgcta	cacttgccag	2340
cgccctagcg	cccgcctcct	tcgctttctt	cccttccttt	ctcgccacgt	tcgcccggctt	2400
tccccgtcaa	gctctaaatc	ggggcatccc	tttagggttc	cgatttagtg	ctttacggca	2460
cctcgacccc	aaaaaacttg	attaggggtg	tggttcacgt	agtgggccat	cgccctgata	2520
gacggttttt	cgccctttga	cgttggagtc	cacgttcttt	aatagtggac	tcttgttcca	2580
aactggaaca	acactcaacc	ctatctcggt	ctattctttt	gatttataag	ggattttggg	2640
gatttcggcc	tattggttaa	aaaatgagct	gatttaacaa	aaatttaacg	cgaattaatt	2700
ctgtggaatg	tgtgtcagtt	aggggtgtgga	aagtccccag	gctccccagg	caggcagaag	2760
tatgcaaagc	atgcatctca	attagtcagc	aaccagggtg	ggaaagtccc	caggctcccc	2820
agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaaccatag	tcccgccect	2880
aactccgccc	atcccgcgcc	taactccgcc	cagttccgcg	cattctccgc	cccatggctg	2940
actaattttt	tttattttatg	cagagggcga	ggcgcctct	gcctctgagc	tattccagaa	3000
gtagttagga	ggcttttttg	gagggcctag	cttttgcaaa	aagctcccgg	gagcttgtat	3060
atccattttc	ggatctgata	aagagacagg	atgaggatcg	tttcgcatga	ttgaacaaga	3120
tggattgcac	gcaggttctc	cggccgcttg	ggtggagagg	ctattcggct	atgactgggc	3180
acaacagaca	atcggctgct	ctgatgccgc	cgtgttccgg	ctgtcagcgc	aggggcgccc	3240
ggttcttttt	gtcaagaccg	acctgtccgg	tgcctgcaat	gaactgcagg	acgaggcagc	3300
gcggctatcg	tggctggcca	cgacggggcg	tccttgccga	gctgtgctcg	acgttgtcac	3360
tgaagcggga	agggactggc	tgctattggg	cgaagtgccg	gggcaggatc	tcctgtcatc	3420
tcaccttgct	cctgcccaga	aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	3480
gcttgatcg	gctacctgcc	cattcgacca	ccaagcgaag	catcgcatcg	agcgagcacg	3540
tactcggatg	gaagcgggtc	tttgcgata	ggatgactcg	gacgaagagc	atcaggggct	3600
cgcgccagcc	gaactggttc	ccaggctcaa	ggcgcgcgat	cccgcggcg	aggatctcgt	3660
cgtgacccat	ggcgatgcct	gcttgccgaa	tatcatgggt	gaaaatggcc	gcttttcttg	3720
attcatcgac	tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	cgttggctac	3780
ccgtgatatt	gctgaagagc	ttggcggcga	atgggctgac	cgcttcctcg	tgctttacgg	3840
tatcgccgct	cccgatccgc	agcgcacatc	cttctatcgc	cttcttgacg	agttcttctg	3900
agcgggactc	tgggggttcgc	gaaatgaccg	accaagcgac	gcccacactg	ccatcacgag	3960
atttcgatcc	caccgcgcgc	ttctatgaaa	ggttgggctt	cggaatcggt	ttccgggacg	4020
cgggttggtg	gatcctccag	cgcgggggat	tcagtctgga	gttcttcgcc	caccccaact	4080
tgtttattgc	agctttataat	ggttacaaat	aaagcaatag	catcacaaat	ttcacaaata	4140
aagcattttt	ttactgcat	tctagttgtg	gtttgtccaa	actcatcaat	gtatcttata	4200
atgtctgtat	accgtcgacc	tctagctaga	gcttggcgta	atcatggtca	tagctgtttc	4260
ctgtgtgaaa	ttgttatccg	ctcacaaatc	cacacaacat	acgagccgga	agcataaagt	4320
gtaaagcctg	gggtgcctaa	tgagtgaagt	aactcacatt	aattgcgttg	cgctcactgc	4380
ccgctttcca	gtcgggaaac	ctgtcgtgcc	agctgcatta	atgaatcggc	caacgcgcgg	4440
ggagaggcgg	tttgcgatatt	gggcgctctt	ccgcttcctc	gctcactgac	tcgctgcgct	4500
cggtcgttcg	gctgcggcga	gcggtatcag	ctcactcaaa	ggcggttaata	cggttatcca	4560
cagaatcagg	ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	4620
accgtaaaaa	ggccgcggtg	ctggcgtttt	tccataggtc	ccgccccctt	gacgagcatc	4680
acaaaaatcg	acgctcaagt	cagagggtggc	gaaacccgac	aggactataa	agataaccagg	4740
cgtttccccc	tggaaagctcc	ctcgtgcgct	ctcctgttcc	gaccctgccg	cttaccggat	4800
acctgtccgc	ctttctccct	tcgggaagcg	tggcgctttc	tcaatgctca	cgctgtaggt	4860
atctcagttc	ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccggttc	4920
agcccgaccg	ctgcgcctta	tccggtaact	atcgtcttga	gtccaacccg	gtaagacacg	4980
acttatcgcc	actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	5040
gtgctacaga	gttcttgaag	tgggtggccta	actacggcta	cactagaagg	acagtatttg	5100
gtatctgcgc	tctgctgaag	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	5160
gcaaacacaa	caccgctggg	agcggtggtt	tttttgtttg	caagcagcag	attacgcgca	5220
gaaaaaaagg	atctcaagaa	gatccttttg	tcttttctac	ggggtctgac	gctcagtgga	5280
acgaaaactc	acgttaaggg	attttgggtc	tgagattatc	aaaaaggatc	ttcacctaga	5340
tccttttaaa	ttaaaaatga	agttttaaat	caatctaaag	tatatatgag	taaacttggt	5400

```

ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcggt 5460
catccatagt tgccctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat 5520
ctggcccccag tgctgcaatg ataccgcgag acccacgctc accgggtcca gatttatcag 5580
caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct 5640
ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt 5700
tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg 5760
cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc atgttgtgca 5820
aaaaagcggg tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt 5880
tatcactcat ggttatggca gcaactgcata attctcttac tgtcatgcc a tccgtaagat 5940
gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt atgcgggcgac 6000
cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa 6060
aagtgtccat cattggaaaa cgcttcttcg ggcgaaaact ctcaaggatc ttaccgctgt 6120
tgagatccag ttcgatgtaa ccactcgtg caccacaactg atcttcagca tcttttactt 6180
tcaccagcgt ttctgggtga gcaaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa 6240
gggcgacacg gaaatgttga atactcatac tcttctttt tcaatattat tgaagcattt 6300
atcagggtta ttgtctcatg agcggataca tatttgatg tatttagaaa aataaacaaa 6360
taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtc 6404

```

<210> 2

<211> 6420

<212> DNA

<213> Artificial Sequence

<220>

<223> vector pVP22/Myc-His-TOPO

<400> 2

```

gacggatcgg gagatctccc gatccccctat ggtcgactct cagtacaatc tgcctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagtatac acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgag atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacgggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgccatt gagtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggg aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg 780
gtaggcgtgt acgggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagt 900
taagcttatt atgacctctc gccgctccgt gaagtcgggt ccgcgggagg ttccgcgcga 960
tgagtacgag gatctgtact acaccccgtc ttcaggtatg gcgagtcctc atagtccgcc 1020
tgacacctcc cgccgtggcg ccctacagac acgctcgcgc cagaggggag aggtccggtt 1080
cgtccagtac gacgagtcgg attatgccct ctacggggggc tcgtcttcgg aagacgacga 1140
acacccggag gtcccccgga cgcggcgctc cgtttccggg gcggttttgt ccggcccggg 1200
gctgcgcgg gcgcctccgc caccgcgtg gtccggaggg gccggacgca caccaccac 1260
cgcccccggg gcccccgaa ccagcgggt ggcgactaag gcccccgcg ccccgcgcc 1320
ggagaccacc cgcggcagga aatcggccca gccagaatcc gccgcactcc cagacgcccc 1380
cgcgtcgacg gcgccaaccc gatccaagac acccgcgag gggctggcca gaaagctgca 1440
ctttagcacc gccccccaa accccgacgc gccatggacc ccccggtgg ccggttttaa 1500
caagcgcgtc ttctgcgccg cggtcgggag cctggcggcc atgcatgccc ggatggcggc 1560
ggtccagctc tgggacatgt cgcgtccgcg cacagacgaa gacctaacc aactccttgg 1620
catcaccacc atccgcgtga cggctctgca gggcaaaaac ctgcttcagc gcgccaacga 1680
gttggtgaat ccagacgtgg tgcaggacgt cgacgcggcc acggcgactc gagggcggtc 1740
tgcgcgctcg cgccccacc agcgacctgc agccccagcc cgtccgctt ctgcgccag 1800
acggcccgctc gagggatacc agctcgatc cactagtcca gtgtggtgga attgccctta 1860
agggaattc tgcgatatac cagcacagt gcggccgctc gagtctagag ggcccgcggt 1920
tcgaacaaaa actcatctca gaagaggatc tgaatatgca taccggtcat catcaccatc 1980
accattgagt ttaaaccgcg tgatcagcct cgactgtgcc ttctagttgc cagccatctg 2040

```



```

tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc 5760
tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 5820
tcccccatgt tgtgcaaaaa agcggtttagc tccttcggtc ctccgatcgt tgtcagaagt 5880
aagttggccg cagtgttata actcatgggt atggcagcac tgcataattc tcttactgtc 5940
atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 6000
tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 6060
catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt ctccggggcg aaaactctca 6120
aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtgcacc caactgatct 6180
tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 6240
gcaaaaaagg gaataagggc gacacggaaa tggtgaatac tcatactctt cctttttcaa 6300
tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 6360
tagaaaaata aacaaatagg ggttccgcgc acatttccc gaaaagtgcc acctgacgtc 6420

```

```

<210> 3
<211> 12
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Peptide linker

```

```

<400> 3
Leu Ala Arg Leu Leu Ala Arg Leu Leu Ala Arg Leu
  1             5             10

```

```

<210> 4
<211> 67
<212> PRT
<213> Artificial sequence

```

```

<220>
<223> Conserved sequence of steroid/thyroid hormone
      receptor superfamily DNA-binding domain

```

```

<221> VARIANT
<222> (0)...(0)
<223> Xaa is non-conserved amino acids within the
      DNA-binding domain

```

```

<221> VARIANT
<222> 7, 9, 11, 13, 22, 27, 58, 61, 66
<223> amino acid residues that are almost universally
      conserved, but for which variations have been
      found in some identified hormone receptors

```

```

<400> 4
Cys Xaa Xaa Cys Xaa Xaa Asp Xaa Ala Xaa Gly Xaa Tyr Xaa Xaa Xaa
  1             5             10             15
Xaa Cys Xaa Xaa Cys Lys Xaa Phe Phe Xaa Arg Xaa Xaa Xaa Xaa Xaa
      20             25             30
Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Lys
      35             40             45
Xaa Xaa Arg Xaa Xaa Cys Xaa Xaa Cys Arg Xaa Xaa Lys Cys Xaa Xaa
      50             55             60
Xaa Gly Met
65

```

```

<210> 5
<211> 34
<212> DNA

```

202070-LEB2E660

<213> Bacteriophage P1

<400> 5

ataacttcgt atagcatata ttatacgaag ttat

34

<210> 6

<211> 34

<212> DNA

<213> E. coli

<400> 6

ataacttcgt atagtataca ttatacgaag ttat

34

<210> 7

<211> 34

<212> DNA

<213> E. coli

<400> 7

acaacttcgt ataattgtatg ctatacgaag ttat

34

<210> 8

<211> 34

<212> DNA

<213> Saccharomyces cerevisiae

<400> 8

gaagttccta ttctctagaa agtataggaa cttc

34

<210> 9

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide linker moiety

<400> 9

Gly Gly Gly Gly Ser

1

5

<210> 10

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety; sequence can be repeated
indefinit number of times

<400> 10

Gly Gly Gly Gly Ser

1

5

<210> 11

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

202070" ZEBZEBD

<400> 11
 Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Ser
 1 5 10

<210> 12
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 12
 Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly
 1 5 10

<210> 13
 <211> 18
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 13
 Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Ser Gly Ser Thr
 1 5 10 15
 Lys Gly

<210> 14
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 14
 Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly
 1 5 10

<210> 15
 <211> 18
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Peptide linker moiety

<400> 15
 Gly Ser Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Thr
 1 5 10 15
 Lys Gly

<210> 16
 <211> 14
 <212> PRT
 <213> Artificial Sequence

2020/07/07 14:00:00

<220>

<223> Peptide linker moiety

<400> 16

Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Glu Phe
1 5 10

<210> 17

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<400> 17

Ser Arg Ser Ser Gly
1 5

<210> 18

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<400> 18

Ser Gly Ser Ser Cys
1 5

<210> 19

<211> 27

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<400> 19

Met Gly Arg Ser Gly Gly Gly Cys Ala Gly Asn Arg Val Gly Ser Ser
1 5 10 15
Leu Ser Cys Gly Gly Leu Asn Leu Gln Ala Met
20 25

<210> 20

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

<221> VARIANT

<222> (0)...(0)

<223> Xaa is (GmS)n, where m is
from 2 to 4 and n is from 1 to 11.

<400> 20

Ala Met Xaa Ala Met

00937837 010702

1

5

<210> 21

<211> 16

<212> PRT

<213> Drosophila acanthoptera

<400> 21

Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys
1 5 10 15

09937837.010702